What Is Claimed Is:

1.   A method of dynamically controlling the rate of communication between two entities, comprising:

        receiving an electronic communication, for a first channel between a first entity and a second entity, at a relay element situated between the first entity and the second entity;

        retrieving from said communication a first value associated with a first target bandwidth for said first channel;

        determining whether said relay element can provide said first target bandwidth for said first channel; and

        modifying said first value in said communication if said relay element cannot provide said first target bandwidth for said channel.

2.   The method of claim 1, further comprising:

        forwarding said communication;

        wherein said first value in said forwarded communication indicates a bandwidth allocated to said first channel by said relay element.

3.   The method of claim 1, further comprising, prior to said determining:

        receiving a set of communications on a set of channels through said switching element, not including said first channel;

        retrieving from said set of communications a set of values associated with target bandwidths for said set of channels; and

        summing said target bandwidths to calculate a total allocated bandwidth for said relay element.

23

4.   The method of claim 3, wherein said determining comprises:

comparing said total allocated bandwidth to a maximum bandwidth of said relay element; and

5   if said maximum bandwidth exceeds said total allocated bandwidth by a difference of more than said first target bandwidth, determining that said relay element can provide said first target bandwidth for said first channel.

5.   The method of claim 1, wherein said determining comprises:

10   comparing said first target bandwidth for said first channel to a previous bandwidth granted to said first channel by said relay element; and

if said first target bandwidth is greater than said previous bandwidth, comparing a difference between said first target bandwidth and said previous bandwidth with an unallocated bandwidth of said relay element.

15

6.   The method of claim 1, wherein said modifying comprises changing said first value to a value associated with zero bandwidth.

7.   The method of claim 1, wherein said communication includes said

20   first value and a second value associated with a requested bandwidth for said first channel; and

wherein said first value is modifiable and said second value is not modifiable.

25   8.   The method of claim 1, wherein said first value is a time value representing a time between communication transmissions from the first entity to the second entity on said first channel.

24

9.   The method of claim 1, wherein said electronic communication is a packet.

5   10.   The method of claim 9, wherein said relay element is a switch and wherein said first entity and said second entity are computer systems.

11.   The method of claim 1, wherein one of said first entity and said second entity is a computer system; and

10   wherein the other of said first entity and said second entity is an input/output subsystem.

12.   A computer readable storage medium storing instructions that, when executed by a computer, cause the computer to perform a method of

15   dynamically controlling the rate of communication between two entities, the method comprising:

          receiving an electronic communication, for a first channel between a first entity and a second entity, at a relay element situated between the first entity and the second entity;

20   retrieving from said communication a first value associated with a first target bandwidth for said first channel;

          determining whether said relay element can provide said first target bandwidth for said first channel; and

          modifying said first value in said communication if said relay element

25   cannot provide said first target bandwidth for said channel.

13.   A method of dynamically controlling the rate of communication

25

between two entities, comprising:

generating at a first entity a first electronic communication for transmission to a second entity over a first communication channel, wherein said first communication includes a first value indicating a target rate of

5 communication for said channel;

receiving said first communication at a switching element;

determining whether said switching element can provide said target rate of communication for said first channel;

if said switching element cannot provide said target rate of

10 communication, altering said first value to indicate a lower target rate of communication for said first channel;

receiving said first communication at said second entity; and

communicating said first value to said first entity.

15   14.   The method of claim 13, further comprising determining whether said switching element previously allocated a rate of communication to said first channel.

     15.   The method of claim 13, further comprising after said

20 communicating:

transmitting one or more communications from said first entity toward said second entity at said lower target rate of communication.

     16.   The method of claim 13, wherein said generating comprises storing

25 said first value in said first communication prior to transmitting it over said first channel.

17.   The method of claim 16, wherein said generating further comprises storing a second value in said first communication; and

        wherein said second value indicates a requested rate of communication for said channel.

5

18.   The method of claim 17, wherein said first value is equal to said second value.

19.   The method of claim 17, wherein one or more of said first value
10   and said second value comprises a threshold value indicating a maximum rate of communication.

20.   The method of claim 17, wherein one or more of said first value and said second value comprise a threshold value indicating a minimum rate of
15   communication.

21.   The method of claim 20, further comprising at said switching element:

        detecting said threshold value indicating said minimum rate of
20   communication; and

        tearing down said channel.

22.   The method of claim 17, wherein one or more of said first value and said second value comprises a time period representing a delay between
25   transmission of successive communications over said first channel from said first entity; and

        wherein said rate of communication indicated by said time period is

27

substantially equal to the inverse of said time period.

23.   The method of claim 13, wherein said determining comprises:

determining whether a maximum rate of communication of said switching
element has been allocated; and

if said maximum rate has not been allocated, identifying an available rate
of communication of said switching element.

24.   The method of claim 23, wherein said identifying comprises:

(a)   receiving a communication prior to said first communication at
said switching element, on a channel other than said first channel;

(b)   allocating a portion of a maximum rate of communication of said
switching element to said other channel;

(c)   repeating said steps (a) - (b);

(d)   summing said rates of communication allocated to said other
channels to determine a total allocated rate of communication; and

(e)   determining the different between said maximum rate of
communication and said total allocated rate of communication.

25.   The method of claim 24, wherein said repeating comprises
repeating steps (a) - (b) for a predetermined period of time.

26.   The method of claim 13, wherein said altering comprises setting
said first value to a threshold value indicating a minimum rate of communication.

27.   The method of claim 26, further comprising at said first entity after
said communicating:

28

ceasing transmission of communications to said second entity over said first channel.

28. The method of claim 13, wherein said first value is a time period between successive electronic communication transmissions from said first entity on said first channel.

29. The method of claim 28, wherein said target rate of communication is substantially equal to the inverse of said first value.

30. The method of claim 13, wherein said first value is a measure of bandwidth.

31. A computer readable storage medium storing instructions that, when executed by a computer, cause the computer to perform a method of dynamically controlling the rate of communication between two entities, the method comprising:

generating at a first entity a first electronic communication for transmission to a second entity over a first communication channel, wherein said first communication includes a first value indicating a target rate of communication for said channel;

receiving said first communication at a switching element;

determining whether said switching element can provide said target rate of communication for said first channel;

if said switching element cannot provide said target rate of communication, altering said first value to indicate a lower target rate of communication for said first channel;

29

receiving said first communication at said second entity; and

communicating said first value to said first entity.

32.   A method of controlling a network communication rate,
5   comprising:

receiving at an intermediate node coupling a first network node and a
second network node a rate value representing a rate of communication between
the first network node and the second network node; and

if the intermediate node cannot conduct communications between the first
10   network node and the second network node at said rate value, decreasing said rate
value such that the intermediate node can conduct communications between the
first network node and the second network node at said rate value.

33.   The method of claim 32, wherein said rate value is a time between
15   communications transmitted from the first network node toward the second
network node.

34.   The method of claim 33, wherein said decreasing comprises
increasing said time between communications.
20

35.   The method of claim 32, wherein if said rate value is decreased to a
first value, the first network node stops sending communications toward the
second network node through the intermediate node.

25   36.   The method of claim 32, wherein if said rate value received at the
intermediate node has a second value, the first network node sends
communications toward the second network node through the intermediate node

30

at a maximum rate.

37.   The method of claim 32, further comprising:

notifying the first network node of said decreased rate value;

wherein the first network node then transmits communications toward the
second network node at said decreased rate value.

38.   The method of claim 32, wherein said rate value is a target rate
value.

39.   The method of claim 38, further comprising:

receiving at the intermediate node from the first network node a requested
rate value representing a requested rate of communication between the first
network node and the second network node.

40.   The method of claim 32, wherein the intermediate node is
InfiniBand compliant.

41.   The method of claim 32, wherein the intermediate node is a switch.

42.   The method of claim 32, wherein the intermediate node is a router.

43.   The method of claim 32, wherein the intermediate node is a hub.

44.   The method of claim 32, wherein the intermediate node is a bridge.

45.   The method of claim 32, wherein the intermediate node is a

31

repeater.

46.   The method of claim 32, wherein the intermediate node is a network adapter.

5

47.   The method of claim 32, wherein the intermediate node is a computer.

48.   The method of claim 32, wherein the intermediate node is a
10   communication bus.

49.   A computer readable storage medium storing instructions that, when executed by a computer, cause the computer to perform a method of controlling a network communication rate, the method comprising:

15   receiving at an intermediate node coupling a first network node and a second network node a rate value representing a rate of communication between the first network node and the second network node; and

if the intermediate node cannot conduct communications between the first network node and the second network node at said rate value, decreasing said rate
20   value such that the intermediate node can conduct communications between the first network node and the second network node at said rate value.

50.   A method of controlling a network traffic rate, comprising:

sending a rate value from a first network node toward a second network
25   node, wherein said rate value represents a rate of traffic between the first network node and the second network node;

at one or more intermediate nodes between the first network node and the

32

second network node:

receiving said rate value;

if the intermediate node cannot communicate traffic between the first network node and the second network node at said rate value,

5   decreasing said rate value to a value at which the intermediate node can communicate traffic between the first network node and the second network node; and

forwarding said rate value toward the second network node;
and

10   communicating between the first network node and the second network node at said rate value.

51.   A computer readable storage medium storing instructions that, when executed by a computer, cause the computer to perform a method of

15   controlling a network traffic rate, the method comprising:

sending a rate value from a first network node toward a second network node, wherein said rate value represents a rate of traffic between the first network node and the second network node;

at one or more intermediate nodes between the first network node and the

20   second network node:

receiving said rate value;

if the intermediate node cannot communicate traffic between the first network node and the second network node at said rate value, decreasing said rate value to a value at which the intermediate node can

25   communicate traffic between the first network node and the second network node; and

forwarding said rate value toward the second network node;

33

and

communicating between the first network node and the second network node at said rate value.

5   52.   A computer readable storage medium containing a data structure configured to indicate a rate of communication over a communication channel, the data structure comprising:

    a header portion comprising:

        an identifier of an originator of said data structure;

10   an identifier of a destination of said data structure; and

        a first value corresponding to a target rate of communication between said originator and said destination;

    wherein said first value is modifiable during transmission of said data structure from said originator to said destination.

15   53.   The computer readable storage medium of claim 52, wherein said first value of said header portion of said data structure comprises a time period and said target rate of communication is substantially equal to the inverse of said time period.

20   54.   The computer readable storage medium of claim 52, said data structure further comprising:

    a data portion comprising a set of data.

25   55.   The computer readable storage medium of claim 52, said header portion of said data structure further comprising:

    a second value corresponding to a requested rate of communication

34

between said originator and said destination.

56.　A network node for dynamically controlling a network rate of communication, comprising:

5　　a communication port configured to conduct communications from a first network node toward a second network node; and

logic coupled to said communication port, wherein said logic is configured to:

identify a rate value representing a rate of communication between

10　the first network node and the second network node; wherein said rate value was originated by the first network node; and

decrease said rate value if the network node cannot conduct communications between the first network node and the second network node at the rate value.

15

57.　An apparatus for dynamically adjusting the rate of communications between a first entity and a second entity on a channel, comprising:

a communication port configured to forward a communication received from a first entity toward a second entity on a communication channel;

20　　a first memory configured to store said communication;

a second memory configured to store a target bandwidth for said channel, wherein said target bandwidth is indicated by a first value in said communication;

a comparator configured to compare said target bandwidth to an available bandwidth for said port; and

25　　a processor configured to adjust said first value to indicate a different target bandwidth.

35

Inventor: Mott

58. The apparatus of claim 57, further comprising an extractor configured to extract said first value from said communication.

59. The apparatus of claim 58, wherein said value comprises a time period representing a delay between communication transmissions from said first entity toward said second entity on said channel, the apparatus further comprising:

an inverter configured to invert said time period;

wherein said target bandwidth is substantially equal to said inverted time period.

60. The apparatus of claim 59, further comprising:

an adder configured to add said target bandwidth of said communication to a target bandwidth of a previous communication on a different channel to calculate a total allocated bandwidth.

61. The apparatus of claim 60, wherein said available bandwidth is substantially equal to a maximum bandwidth of said port minus said total allocated bandwidth.

62. The apparatus of claim 58, wherein said extractor is further configured to retrieve a second value from said communication;

wherein said second value indicates a requested bandwidth for said channel.

63. The apparatus of claim 57, wherein said processor is configured to adjust said first value to indicate a lower target bandwidth if said apparatus is unable to provide said target bandwidth.

36

64.   A communication system configured for dynamic rate flow control

between two communicating devices, comprising:

a first device configured to generate a communication for transmission

5   toward a second device over a first channel, wherein said communication includes

a first value indicating a target bandwidth for said first channel;

a switch element configured to receive said communication and direct said

communication toward said second device, wherein said switch element alters

said first value if said switch element cannot provide said target bandwidth for

10   said first channel; and

a second device configured to receive said communication and report said

first value to said first entity.

37

Inventor: Mott